Proceedings of the Royal Society of Edinburgh, **116A**, 341–358, *1990*

On the equation grad $f = M$ grad g

Max Jodeit, Jr. and **Peter J. Olver***
School of Mathematics, University of Minnesota, Minneapolis, MN 55455, U.S.A.

(MS received 18 January 1990. Revised MS received 12 April 1990)

Synopsis

The system of differential equations $\nabla f = M\nabla g$, where M is a given square matrix, arises in many contexts. A complete solution to this problem in the case when M is a constant matrix is presented here. Applications to continuum mechanics and biHamiltonian systems are indicated.

1. Introduction

In this paper we consider the elementary system of partial differential equations

$$\nabla f = M\nabla g, \tag{1.1}$$

in which $f(x)$ and $g(x)$ are scalar-valued functions on an open domain in an n-dimensional vector space, either real or complex, and $M(x)$ is a given $n \times n$ matrix of functions. Thus, (1.1) constitutes a system of n linear partial differential equations for the functions f, g, and we are interested in its general solution. The basic equation (1.1) arises in many different contexts, including the study of conservation laws in elasticity [**4**, **6**], phase transitions [**1**], and biHamiltonian systems [**7**]. Despite its seeming simplicity, the general equation (1.1) is not well understood, and we are unaware of any systematic treatments in the literature. In this paper we shall exclusively consider the case when M is a constant matrix: our main result is a complete classification of all solutions in this case. Except for some interesting special cases, cf. [**7**], the case of non-constant matrices M of size larger than 2×2 is, to our knowledge, unstudied and completely open.

As a first step, we note that, when M is constant, the integrability conditions for (1.1) are the system of second order equations

$$M\nabla^2 g = \nabla^2 g M^T. \tag{1.2}$$

The goal then is to describe the general solution to this system. We shall be primarily interested in (real or complex) analytic solutions to this system, although we shall make some remarks on less differentiable solutions in the final section, which also includes the motivating problem from Ball and James' work on phase transitions. The basic line of attack is to perform a linear change of variables so as to place the matrix M in as simple a form as possible. Because the linear transformation $\bar{x} = Ax$ transforms M into the similar matrix $\bar{M} = A^{-T}MA^T$, we can choose coordinates so that M is in Jordan canonical form, or, in the real case, real normal form. Then, to simplify matters, we work our way up to the

* Research supported in part by NSF Grant DMS 89-01600.

most general case, finding the general solution to (1.1) in the cases when:

(i) M consists of a single Jordan block – the irreducible case;

(ii) M has only one eigenvalue;

(iii) M is a general complex matrix in Jordan canonical form; and, finally,

(iv) M is in real normal form.

2. The irreducible case

We begin our study of (1.1) by assuming that M is an *irreducible* complex matrix, which means that it has only one eigenvalue, and its Jordan canonical form consists of a single Jordan block. We are interested in complex analytic solutions $f(x)$, $g(x)$. Choosing an appropriate basis of our underlying complex vector space V, we have

$$M = \lambda I + U, \tag{2.1}$$

where λ is the eigenvalue, I is the $(n+1)\times(n+1)$ identity matrix, and U is the $(n+1)\times(n+1)$ upper triangular matrix with 1's on the super-diagonal and 0's elsewhere, so n equals the number of 1's in U. Vectors in V are written as

$$\mathbf{x} = [x_0, x_1, x_2, \ldots, x_n]^T$$

relative to the Jordan basis of M. (All vectors are column vectors, the superscript ^{T} denoting transpose.) We call x_0 the *major variable* for the Jordan block; the other variables $x_1, \ldots, x_n$ are called *minor variables*. Let t be a scalar parameter. Define the scalar variable

$$\xi(t) = x_0 + tx_1 + t^2x_2 + \ldots + t^nx_n, \tag{2.2}$$

corresponding to the single Jordan block M.

THEOREM 2.1. *Let $f(\mathbf{x})$, $g(\mathbf{x})$ be analytic functions on a convex domain in the complex vector space V. Then f, g satisfy* (1.1) *where M is a single Jordan block* (2.1) *if and only if there exist smooth scalar-valued functions $a_0(s), a_1(s), \ldots, a_n(s)$ and a constant c such that*

$$f(\mathbf{x}) = f_0(\mathbf{x}) + f_1(\mathbf{x}) + \ldots + f_n(\mathbf{x}) + c, \quad g(\mathbf{x}) = g_0(\mathbf{x}) + g_1(\mathbf{x}) + \ldots + g_n(\mathbf{x}),$$

where

$$f_k(\mathbf{x}) = \lambda \frac{\partial^k}{\partial t^k} a_k(\xi(t))\Big|_{t=0} + k\frac{\partial^{k-1}}{\partial t^{k-1}} a_k(\xi(t))\Big|_{t=0}, \tag{2.3}$$

$$g_k(\mathbf{x}) = \frac{\partial^k}{\partial t^k} a_k(\xi(t))\Big|_{t=0}, \tag{2.4}$$

where $\xi(t)$ is given by (2.2).

EXAMPLE 2.2. We illustrate what these formulae look like for small values of k.

$k=0$:
$$\begin{aligned} g_0(\mathbf{x}) &= a_0(x_0 + tx_1 + t^2x_2 + \ldots)|_{t=0} = a_0(x_0), \\ f_0(\mathbf{x}) &= \lambda g_0(\mathbf{x}) = \lambda a_0(x_0). \end{aligned} \tag{2.5-0}$$

Thus any function of the major variable x_0 provides a solution to (1.1), (2.1).

$k=1$:

$$\begin{aligned} g_1(\mathbf{x}) &= \frac{\partial}{\partial t} a_1(x_0 + tx_1 + t^2x_2 + \dots)\big|_{t=0} \\ &= x_1 a_1'(x_0) \\ f_1(\mathbf{x}) &= \lambda g_1(\mathbf{x}) + a_1(x_0 + tx_1 + t^2x_2 + \dots)\big|_{t=0} \\ &= \lambda x_1 a_1'(x_0) + a_1(x_0) \end{aligned} \tag{2.5-1}$$

$k=2$:

$$\begin{aligned} g_2(\mathbf{x}) &= \frac{\partial^2}{\partial t^2} a_2(x_0 + tx_1 + t^2x_2 + \dots)\big|_{t=0} \\ &= x_1^2 a_2''(x_0) + 2x_2 a_2'(x_0) \\ f_2(\mathbf{x}) &= \lambda g_2(\mathbf{x}) + 2\frac{\partial}{\partial t} a_2(x_0 + tx_1 + t^2x_2 + \dots)\big|_{t=0} \\ &= \lambda x_1^2 a_2''(x_0) + 2\lambda x_2 a_2'(x_0) + 2x_1 a_2'(x_0). \end{aligned} \tag{2.5-2}$$

$k=3$:

$$\begin{aligned} g_3(\mathbf{x}) &= \frac{\partial^3}{\partial t^3} a_3(x_0 + tx_1 + t^2x_2 + \dots)\big|_{t=0} \\ &= x_1^3 a_3'''(x_0) + 6x_1x_2 a_3''(x_0) + 6x_3 a_3'(x_0) \\ f_3(\mathbf{x}) &= \lambda g_3(\mathbf{x}) + 3\frac{\partial^2}{\partial t^2} a_3(x_0 + tx_1 + t^2x_2 + \dots)\big|_{t=0} \\ &= \lambda g_3(\mathbf{x}) + 3x_1^2 a_3''(x_0) + 6x_2 a_3'(x_0). \end{aligned} \tag{2.5-3}$$

Thus, according to Theorem 2.1, the general solution to (1.1) when M is a 4×4 Jordan block with eigenvalue λ is

$$\begin{aligned} f(\mathbf{x}) &= \lambda g(\mathbf{x}) + 3x_1^2 a_3''(x_0) + 6x_2 a_3'(x_0) + 2x_1 a_2'(x_0) + a_1(x_0) + c, \\ g(\mathbf{x}) &= x_1^3 a_3'''(x_0) + 6x_1x_2 a_3''(x_0) + 6x_3 a_3'(x_0) + x_1^2 a_2''(x_0) \\ &\quad + 2x_2 a_2'(x_0) + x_1 a_1'(x_0) + a_0(x_0), \end{aligned}$$

where c is a constant, and a_0, a_1, a_2, a_3 are arbitrary analytic scalar functions of the major variable x_0. Note that in particular the solution is necessarily a polynomial function of the minor variables.

More generally, it can easily be seen that the solutions (2.3), (2.4), of (1.1), (2.1) must be certain polynomial functions of the minor variables, whose coefficients are derivatives of arbitrary functions of the major variable. It can be shown that these polynomials are well known in combinatorial theory. Indeed, the solutions (2.3), (2.4), have the explicit form

$$\begin{aligned} g_k(\mathbf{x}) &= \sum_{j=1}^{k} P_{k,j}(x_1, x_2, \dots, x_{k-j+1}) a_k^{(j)}(x_0), \\ f_k(\mathbf{x}) &= \lambda g_k(\mathbf{x}) + k \sum_{j=1}^{k-1} P_{k-1,j}(x_1, x_2, \dots, x_{k-j}) a_k^{(j)}(x_0), \end{aligned}$$

where $a_k^{(j)}$ is the jth derivative of a_k. The polynomials $P_{k,j}$ are essentially the *partial Bell polynomials* which arise in the Faa-di-Bruno formula for the derivatives of the composition of two functions; using the notation in Comtet [**2**], we find that

$$P_{k,j}(x_1, x_2, \ldots, x_j) = B_{k,j}(x_1, 2x_2, 6x_3, \ldots, j!\, x_j).$$

See **[2]** for a table of these polynomials for $k \leqq 12$.

We remark that any solution of (1.1) for an $(n+1) \times (n+1)$ Jordan block with eigenvalue λ is obviously also a solution for any larger sized Jordan block having the same eigenvalue. Correspondingly, if we have a solution f, g which does not explicitly depend on the highest order minor variable x_n, then it is also a solution to the same problem for the smaller $n \times n$ Jordan block with the same eigenvalue. For instance, in Example 2.2, if we set $a_3 = 0$, then we recover the general solution for the 3×3 Jordan block. This remark will be the key to our inductive proof of the theorem. Conceptually, it is often helpful to replace the parametrised variable $\xi(t)$, cf. (2.2), by a formally infinite power series

$$\xi(t) = x_0 + tx_1 + t^2x_2 + \ldots, \tag{2.6}$$

corresponding to an "infinite Jordan block". This does not change any of the formulae (2.3), (2.4), and avoids having to keep track of the precise number of terms in ξ at each stage.

Proof of Theorem 2.1.

In outline, the proof of this result, and our subsequent more complicated versions, always proceeds in two steps. We first demonstrate, by direct computation, that the explicit formulae (2.3), (2.4) always give solutions to (1.1), (2.1). Then, to demonstrate that these are the only solutions, we use a simple reverse induction on n, the size of the Jordan block. Indeed, by a direct analysis we determine the dependence of a general solution to our problem on the highest order minor variable x_n. Convexity of the underlying domain will imply that we can find a particular solution of the form (2.3), (2.4) with the same highest order terms. We then use the linearity of (1.1) to subtract off our particular solution, resulting in a solution to (1.1), (2.1) which does not depend on x_n, which, by the above remark, is then a solution to the problem for the next smaller size Jordan block. Induction will then complete the proof.

We can further simplify matters by defining

$$h(\mathbf{x}) = f(\mathbf{x}) - \lambda g(\mathbf{x}). \tag{2.7}$$

Then h, g satisfy the system

$$\nabla h = U \nabla g, \tag{2.8}$$

which is just (1.1) in the case when M is a single Jordan block with eigenvalue 0. Written out in detail, (2.8) reads

$$\frac{\partial h}{\partial x_j} = \frac{\partial g}{\partial x_{j+1}}, \quad j = 0, \ldots, n-1, \tag{2.9}$$

$$\frac{\partial h}{\partial x_n} = 0. \tag{2.10}$$

The following elementary lemma is the key to proving the first part of the theorem.

LEMMA 2.3. *Let $a(s)$ be any smooth scalar function, let $\xi(t)$ be defined by* (2.6), *and let*

$$g_k(\mathbf{x}) = \frac{\partial^k}{\partial t^k} a(\xi(t))\bigg|_{t=0}. \tag{2.11}$$

Then

$$\frac{\partial g_k}{\partial x_j} = \begin{cases} \dfrac{k!}{(k-j)!} \dfrac{\partial^{k-j}}{\partial t^{k-j}} a'(\xi(t))\bigg|_{t=0}, & j \leqq k, \\ 0, & j > k, \end{cases} \tag{2.12}$$

where $a'(s)$ denotes the derivative of $a(s)$.

Thus, if g_k is given by (2.11), and we set

$$h_k(\mathbf{x}) = k \frac{\partial^{k-1}}{\partial t^{k-1}} a(\xi(t))\bigg|_{t=0}, \tag{2.13}$$

then (2.12) and the corresponding formula for h_k immediately proves that g_k, h_k solve (2.9). Moreover, h_k will also satisfy the final condition (2.10) provided $k \leqq n$. Thus, recalling (2.7), and setting $a = a_k$, we have verified the solution (2.4), (2.5).

We now need to prove that there are no other solutions. To this end, we work by induction on n, the scalar case $n = 0$ being trivial. Let g, h be an arbitrary solution to (2.8). By (2.10) h does not depend on x_n. Differentiating (2.9) with respect to x_n, we deduce that $\partial^2 g/\partial x_j \, \partial x_n = 0$ for all $j > 1$ on the domain of definition of g. Since our domain is convex, then this implies that g must have the form

$$g(\mathbf{x}) = b(x_0)x_n + \hat{g}(x_0, x_1, \ldots, x_{n-1}),$$

where b depends only on the major variable. Define the function $a_n(s)$ to be any first integral of $b(s)$, so that $a_n'(s) = b(s)$, and let $g_n(\mathbf{x})$, $h_n(\mathbf{x})$, be the functions given by (2.11), (2.13) with $k = n$, $a = a_n$. Note that, according to (2.12),

$$g_n(\mathbf{x}) = a_n'(x_0)x_n + \hat{g}_n(x_0, x_1, \ldots, x_{n-1}) = b(x_0)x_n + \hat{g}_n(x_0, x_1, \ldots, x_{n-1}),$$

as desired. By linearity, the functions $\tilde{g} = g - g_n$, $\tilde{h} = h - h_n$ satisfy (2.8), and, moreover, depend only on $x_0, x_1, \ldots, x_{n-1}$. Thus, by the above remark, $\tilde{g}$, $\tilde{h}$ must also be a solution to (2.9), (2.10) with n replaced by $n - 1$. Now we can use our inductive hypothesis to complete the proof.

Finally, we remark that, as a consequence of the deRham Theorem **[11]**, we can readily extend Theorem 2.1 to any simply connected domain Ω, although the precise statement is a little more tricky. The functions $a_k(x)$ will satisfy $\partial a_k/\partial x_j = 0$ for $j = 1, \ldots, n$. This means that locally they will still be functions of the major variable x_0 alone; however, globally this may not be the case. For instance, if, over a point c the subset $\Omega_c = \{x \in \Omega : x_0 = c\}$ consists of several disconnected pieces, the function a_k, while constant on each connected component of Ω_c, may attain different values on different components.

3. The reducible case – one eigenvalue

We now turn to the case when M, complex, still has only one eigenvalue, but there are several Jordan blocks. This means that there is a direct sum decomposition of the underlying vector space into a sum of invariant *Jordan subspaces*

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_m,$$

where

$$\dim V_i = n_i + 1,$$

n_i indicating the number of off-diagonal 1's in the corresponding Jordan block of M. We can also assume that we have arranged the Jordan subspaces in decreasing size:

$$n^* = n_1 \geqq n_2 \geqq n_3 \geqq \ldots \geqq n_m > 0,$$

and we use $n^* = \max\{n_i\}$ to indicate the size of the largest Jordan block in M. Relative to this decomposition, the matrix M then has Jordan form

$$M = \operatorname{diag}[M_1, M_2, \ldots, M_m], \quad M_i = \lambda I_i + U_i, \tag{3.1}$$

where, for each i, I_i denotes the $(n_i + 1) \times (n_i + 1)$ identity matrix, and U_i is the upper triangular matrix of the same size with 1's on the super-diagonal and 0's elsewhere. Each vector $\mathbf{x} \in V$ can be written as

$$\mathbf{x} = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^m]^T, \quad \text{where} \quad \mathbf{x}^i = [x_0^i, x_1^i, \ldots, x_{n_i}^i]^T.$$

We call the variable x_0^i the *major variable* for the subspace V_i; the other variables are called *minor variables*. Note that a diagonalisable matrix does not have any minor variables.

As in the irreducible case, for each subspace V_i we define the parametrised scalar variable

$$\xi^i(t) = x_0^i + tx_1^i + t^2x_2^i + \ldots + t^{n_i}x_{n_i}^i. \tag{3.2}$$

Further, for each $0 \leqq k \leqq n^*$, set

$$\mu_k = \max\{i \mid n_i \geqq k\},$$

so that μ_k is the number of Jordan blocks in M of size at least $k + 1$, which, by our ordering hypothesis, are the first μ_k blocks. Note that $\mu_0 = m$ is the total number of Jordan blocks, and $\mu_k = 0$ for $k > n^*$. For each $k \leqq n^*$, we define a μ_k-dimensional vector of parametrised variables:

$$\xi^{(k)}(t) = [\xi^1(t), \xi^2(t), \ldots, \xi^{\mu_k}(t)]. \tag{3.3}$$

THEOREM 3.1. *Let $f(\mathbf{x})$, $g(\mathbf{x})$ be analytic functions on a convex domain in V. Then f, g satisfy* (1.1) *where M has only one eigenvalue, and is in Jordan block form* (3.1) *if and only if there exist smooth scalar-valued functions*

$$a_k(s_1, \ldots, s_{\mu_k}), \quad k = 0, \ldots, n^*, \tag{3.4}$$

such that

$$f(\mathbf{x}) = f_0(\mathbf{x}) + f_1(\mathbf{x}) + \ldots + f_{n^*}(\mathbf{x}) + c, \quad g(\mathbf{x}) = g_0(\mathbf{x}) + g_1(\mathbf{x}) + \ldots + g_{n^*}(\mathbf{x}),$$

where c is a constant, and

$$f_k(\mathbf{x}) = \lambda \frac{\partial^k}{\partial t^k} a_k(\xi^{(k)}(t))\bigg|_{t=0} + k \frac{\partial^{k-1}}{\partial t^{k-1}} a_k(\xi^{(k)}(t))\bigg|_{t=0}, \tag{3.5}$$

$$g_k(\mathbf{x}) = \frac{\partial^k}{\partial t^k} a_k(\xi^{(k)}(t))\bigg|_{t=0}, \tag{3.6}$$

where $\xi^{(k)}(t)$ *is given by* (3.2), (3.3).

Proof. The proof is similar to the single Jordan block case. Define $h(\mathbf{x}) = f(\mathbf{x}) - \lambda g(\mathbf{x})$, so (1.1), (3.1) become

$$\frac{\partial h}{\partial x^i_j} = \frac{\partial g}{\partial x^i_{j+1}}, \quad j = 0, \ldots, n_i - 1, \tag{3.7}$$

$$\frac{\partial h}{\partial x^i_{n_i}} = 0. \tag{3.8}$$

Now, if a_k is given by (3.4), and g_k by (3.6), then a chain rule computation as in Lemma 2.3 shows that

$$\frac{\partial g_k}{\partial x^i_j} = \begin{cases} \dfrac{k!}{(k-j)!} \dfrac{\partial^{k-j}}{\partial t^{k-j}} \dfrac{\partial a_k}{\partial s_i}(\xi^{(k)}(t))\bigg|_{t=0}, & j \leqq k \leqq n_i \\ 0, & \text{otherwise.} \end{cases} \tag{3.9}$$

There is a similar formula for the derivatives of h, and it is easy to verify that f, g, as given by (3.5), (3.6), solve (1.1), (3.1).

To check that there are no other solutions, we again work inductively, this time using n^*, the maximal size Jordan block, as our inductive integer. The case $n^* = 0$ is trivial, being the case of a diagonalisable matrix, so $M = \lambda I$; here (1.1) reduces to $\nabla f = \lambda \nabla g$, so $f = \lambda g + c$ for some constant c. In general, the solutions f_k, g_k given by (3.5), (3.6) only depend on x^i_j for $j \leqq k \leqq n_i$, $i = 1, \ldots, \mu_k$; therefore, f_k, g_k also solve (1.1), (3.1) for any M whose first μ_k Jordan blocks (arranged in order of their size) have size at least k. Now suppose g, h are any solution to (3.7), (3.8). Suppose $n^* = n_1 = \ldots = n_r > n_{r+1}$. According to (3.8), h does not depend on the "top order" minor variables $x^1_{n^*}, \ldots, x^r_{n^*}$. Differentiating (3.7) with respect to these variables, and using the convexity of the underlying domain, we conclude that

$$g(\mathbf{x}) = \sum_{i=1}^{r} b_i(x^1_0, \ldots, x^m_0) x^i_{n^*} + \hat{g}(\hat{\mathbf{x}}). \tag{3.10}$$

is an affine function of the top order minor variables, with coefficients depending on all the major variables. Here $\hat{\mathbf{x}}$ denotes all the "lower order" variables x^i_j with $j < n^*$.

The integrability conditions for (3.7), (3.8) are given by (1.2), which, written out in components, reads

$$\frac{\partial^2 g}{\partial x^i_j \, \partial x^k_l} = \frac{\partial^2 g}{\partial x^i_{j-1} \, \partial x^k_{l+1}}, \quad j \geqq 1, \quad l < n_k,$$

$$\frac{\partial^2 g}{\partial x^i_j \, \partial x^k_l} = 0, \qquad j \geqq 1, \quad l = n_k.$$

Consequently, using a simple induction, we find

$$\frac{\partial b_i}{\partial x_0^k} = \frac{\partial^2 g}{\partial x_{n^*}^i \, \partial x_0^k} = \frac{\partial^2 g}{\partial x_0^i \, \partial x_{n^*}^k} = \frac{\partial b_k}{\partial x_0^i} \quad \text{whenever} \quad i, k \leqq r, \tag{3.11}$$

while

$$\frac{\partial b_i}{\partial x_0^k} = \frac{\partial^2 g}{\partial x_{n^*}^i \, \partial x_0^k} = 0 \quad \text{whenever} \quad i \leqq r, \ k > r.$$

This means that the functions $b_1, \ldots, b_r$ can only depend on the "top order major variables" $(x_0^1, \ldots, x_0^r)$, and, moreover, are the coefficients of a closed one-form

$$\omega = b_1(x_0^1, \ldots, x_0^r)\, dx_0^1 + \ldots + b_r(x_0^1, \ldots, x_0^r)\, dx_0^r$$

in these variables. Here (3.11) is equivalent to the closure condition $d\omega = 0$. According to the Poincaré lemma [**5**, **11**], since our domain is convex, there is a function $a_{n^*}(x_0^1, \ldots, x_0^r)$ whose differential is $\omega = da_{n^*}$; therefore

$$\frac{\partial a_{n^*}}{\partial x_0^i} = b_i(x_0^1, \ldots, x_0^r).$$

Now we can match the top order terms of our general solution (3.10) with those of one of our known solutions. Let $g_{n^*}(\mathbf{x})$, $h_{n^*}(\mathbf{x}) = f_{n^*}(x) - \lambda g_{n^*}(\mathbf{x})$, be the corresponding solution given (3.5), (3.6) with $k = n^*$. It is easy to see that

$$g_{n^*}(\mathbf{x}) = \sum_{i=1}^{r} b_i(x_0^1, \ldots, x_0^r) x_{n^*}^i + \hat{g}_{n^*}(\hat{\mathbf{x}})$$

has the same leading order terms as g. Therefore, the differences $\tilde{g} = g - g_{n^*}$, $\tilde{h} = h - h_n^*$ satisfy (3.7), and, moreover, depend only on $\hat{\mathbf{x}}$. Thus, by the above remark, $\tilde{g}$, $\tilde{h}$ must also be a solution to (1.1), but where M has been replaced by a matrix $\tilde{M}$ obtained from M by deleting the last row and column of each maximal sized Jordan block in M. Therefore, $\tilde{M}$ will have Jordan blocks of sizes $n_1 - 1 = \ldots = n_r - 1 \geqq n_{r+1} \geqq \ldots \geqq n_m > 0$. In particular, the maximal sized Jordan block of $\tilde{M}$ has size $n^* - 1$, and we can use our inductive hypothesis to complete the proof. □

4. General complex matrices

Next we treat the case when M is an arbitrary complex matrix. Let M have eigenvalues $\lambda_1, \ldots, \lambda_p$. Break the underlying vector space into a sum of invariant subspaces

$$V = V^1 \oplus V^2 \oplus \ldots \oplus V^p, \tag{4.1}$$

where the generalised eigenspace V^k corresponds to the eigenvalue λ_k,

$$V^k = \ker (M - \lambda_k I)^N, \quad N \gg 0.$$

Each V^k in turn is the direct sum of irreducible invariant subspaces V_i^k corresponding to the distinct Jordan blocks of M. Introduce a basis of V so that M is in Jordan form

$$M = \operatorname{diag}[M^1, M^2, \ldots, M^p],$$

where each submatrix M^k is a $q^k \times q^k$ matrix with just one eigenvalue λ_k, and where q^k is the multiplicity of the eigenvalue λ_k. Each M^k can in turn be expressed as

$$M^k = \operatorname{diag}[M_1^k, M_2^k, \ldots, M_{m_k}^k],$$

where each M_i^k is a single Jordan block of size $(n_i^k + 1) \times (n_i^k + 1)$, with the sizes arranged in decreasing order: $n_1^k \geqq n_2^k \geqq n_3^k \geqq \ldots$. Note that M^k gives the restriction of M to the generalised eigenspace V^k, and M_i^k its restriction to the irreducible subspace V_i^k.

Relative to this basis, the vectors $\mathbf{x}$ in V are written as $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$. It is helpful to introduce some terminology to distinguish which of the Jordan blocks and sub-blocks the various indices on the variables x_i correspond to. The sequence of row numbers corresponding to a large block M^k will be called a *macrointerval*; the corresponding smaller block M_i^k will determine a *microinterval*. Thus, each eigenvalue determines a single macrointerval, which can consist of several smaller microintervals. We shall also describe the variables x_i corresponding to our Jordan basis as belonging to such intervals when their index i belongs to the interval. We let $\mathbf{x}^{(k)}$ denote the column vector of variables x_i belonging to the kth macrointerval, so $\mathbf{x} = [\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}]^T$ corresponds to the decomposition (4.1) of V into generalised eigenspaces. This pattern of macro- and microintervals breaks any square matrix of the same size as M up into large blocks, which we call *macroblocks*, each having size $q^k \times q^l$, each of which in turn is broken up into smaller *microblocks*, corresponding to the M_i^k, each of size $n_i^k \times n_j^l$.

EXAMPLE 4.1. If

$$M = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

so that M has two eigenvalues, 2 and 5, the first corresponding to four Jordan blocks of sizes 2, 2, 1, 1 and the second to a single Jordan block of size 3, the first macrointerval consists of the indices $(1, 2, 3, 4, 5, 6)$ and the second of the remaining indices $(7, 8, 9)$. The microintervals have indices $(1, 2)$, $(3, 4)$, (5), (6), $(7, 8, 9)$, respectively. Thus, relative to this basis, x_3 belongs to the first macrointerval and the second microinterval.

The key lemma that allows us to reduce this general case to the case of a single eigenvalue comes from the general structure of the (symmetric) solutions to the matrix equation $MX = XM^T$, satisfied by $\nabla^2 g$, which follows from the general result on the solution to the matrix equation $AX = XB$ discussed in [**10**, p. 148].

LEMMA 4.2. *The off-diagonal macroblocks of $\nabla^2 g$ are all zero.*

Using this result, and the convexity of our underlying domain, we deduce that the function g breaks up into a sum of functions, each depending on just the variables in a single macrointerval. It is not hard to see that this implies that f must also break up in the same manner. We can therefore apply Theorem 3.1 to each summand, which now is a solution to (1.1) in the case when M has a single eigenvalue to conclude the general solution to our problem in the complex case.

THEOREM 4.3. *For complex matrices M, the general solution to* (1.1) *is given as a sum of solutions to the corresponding single eigenvalue problems*:

$$\begin{aligned} f(\mathbf{x}) &= f^{(1)}(\mathbf{x}^{(1)}) + f^{(2)}(\mathbf{x}^{(2)}) + \ldots + f^{(p)}(\mathbf{x}^{(p)}), \\ g(\mathbf{x}) &= g^{(1)}(\mathbf{x}^{(1)}) + g^{(2)}(\mathbf{x}^{(2)}) + \ldots + g^{(p)}(\mathbf{x}^{(p)}). \end{aligned} \tag{4.2}$$

Here $\mathbf{x}^{(k)}$ denotes the variables in $\mathbf{x}$ corresponding to the kth *macrointerval, i.e. to the* kth *generalised eigenspace, and $f^{(k)}(\mathbf{x}^{(k)})$, $g^{(k)}(\mathbf{x}^{(k)})$, is a solution to the same problem for the submatrix M^k. This latter matrix has just one eigenvalue, and therefore the formulae for $f^{(k)}(\mathbf{x}^{(k)})$, $g^{(k)}(\mathbf{x}^{(k)})$ are given explicitly in Theorem* 3.1.

5. Real matrices

We now turn to the case when the matrix M is real, and we are interested in real solutions to (1.1). We are just allowed to perform real linear changes of coordinates, so we can only reduce M to real normal form. The Jordan blocks corresponding to real eigenvalues are the same as before. Therefore, the statements and proofs of Theorems 2.1 and 3.1 hold without change. In fact, we can relax our smoothness hypothesis on f and g, and assume that f, g are C^∞ functions for the theorem to go through without change. Even more, if a_k is a C^{k+1} function, then (3.5), (3.6), determine a solution to (1.1) which is only C^1 in the major variables, but is a polynomial, and hence analytic, in all the minor variables. This indicates that it is possible to relax the differentiability hypothesis still further, a question we deal with in Section 7.

Turning to the complex eigenvalues, let us begin with the irreducible case. For any complex conjugate pair of eigenvalues $\alpha \pm i\beta$, $\beta \neq 0$, we have irreducible $2n \times 2n$ real Jordan blocks of the form

$$\begin{pmatrix} \Lambda & I & & & \\ & \Lambda & I & & \\ & & \Lambda & I & \\ & & & \Lambda & \cdot \\ & & & & \cdot \quad \cdot \\ & & & & & \cdot \end{pmatrix}, \tag{5.1}$$

where I is the 2×2 identity matrix, and

$$\Lambda = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

is the canonical 2×2 with eigenvalues $\alpha \pm i\beta$. Relative to the Jordan basis of V, the variables will be written as $x_0, y_0, x_1, y_1, \ldots, x_n, y_n$. Let $z_j = x_j + iy_j$, $\bar{z}_j = x_j - iy_j$, and introduce the associated complex derivatives

$$\frac{\partial}{\partial z_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_j} - \frac{\partial}{\partial y_j}\right), \quad \frac{\partial}{\partial \bar{z}_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_j} + \frac{\partial}{\partial y_j}\right).$$

It is easily seen that the real system (1.1) in which M has the form (5.1) is equivalent to the complex system of differential equations

$$\frac{\partial f}{\partial z_j} = (\alpha + i\beta)\frac{\partial g}{\partial z_j} + \frac{\partial g}{\partial z_{j+1}}, \quad j = 0, \ldots, n-1, \quad \frac{\partial f}{\partial z_n} = (\alpha + i\beta)\frac{\partial g}{\partial z_n},$$

$$\frac{\partial f}{\partial \bar{z}_j} = (\alpha - i\beta)\frac{\partial g}{\partial \bar{z}_j} + \frac{\partial g}{\partial \bar{z}_{j+1}}, \quad j = 0, \ldots, n-1, \quad \frac{\partial f}{\partial \bar{z}_n} = (\alpha - i\beta)\frac{\partial g}{\partial \bar{z}_n}.$$

From this, a straightforward calculation shows that the system (1.1), (5.1) can be placed into a convenient purely complex form.

LEMMA 5.1. *Given real functions f, g, define the complex function*

$$F = f - (\alpha - i\beta)g, \tag{5.2}$$

which depends on the complex variables $z_j = x_j + iy_j$, $\bar{z}_j = x_j - iy_j$, $j = 0, \ldots, n$. Then f, g satisfy (1.1), (5.1) *if and only if F is a solution to the system*

$$\frac{\partial F}{\partial \bar{z}_j} = -\sum_{l=1}^{n-j}\left(\frac{-i}{2\beta}\right)^l \frac{\partial \bar{F}}{\partial \bar{z}_{j+l}}, \quad j = 0, \ldots, n. \tag{5.3}$$

THEOREM 5.2. *Let $F(\mathbf{z}, \bar{\mathbf{z}})$ be a complex C^2 function defined on a convex domain in V. Define the parametrised complex variable*

$$\zeta(t) = z_0 + tz_1 + t^2 z_2 + \ldots + t^n z_n. \tag{5.4}$$

Then F satisfies the system of partial differential equations (5.3) *if and only if there exist complex analytic scalar functions $a_0(s), a_1(s), \ldots, a_n(s)$ such that*

$$F(\mathbf{z}, \bar{\mathbf{z}}) = F_0(\mathbf{z}, \bar{\mathbf{z}}) + F_1(\mathbf{z}, \bar{\mathbf{z}}) + \ldots + F_n(\mathbf{z}, \bar{\mathbf{z}}),$$

where

$$F_k(\mathbf{z}, \bar{\mathbf{z}}) = \left\{\frac{\partial^k}{\partial t^k} a_k(\zeta(t)) - \sum_{l=1}^{n-j}\left(\frac{-i}{2\beta}\right)^l \frac{k!}{(k-l)!}\,\overline{\frac{\partial^{k-l}}{\partial t^{k-l}} a_k(\zeta(t))}\right\}\Bigg|_{t=0}. \tag{5.5}$$

Proof. The basic method of proof is the same as always. We begin by showing that (5.5) really does determine a solution to (5.3). Note that the first term on the right-hand side of (5.5) is an analytic function of $z_0, \ldots, z_k$, whereas the summation terms are analytic in the complex conjugate variables $\bar{z}_0, \ldots, \bar{z}_k$

Therefore, using Lemma 2.3, we find

$$\frac{\partial F}{\partial \bar{z}_j} = -\sum_{l=1}^{k-j} \left(\frac{-i}{2\beta}\right)^l \frac{k!}{(k-l-j)!} \left. \overline{\frac{\partial^{k-l-j}}{\partial t^{k-l-j}} a_k'(\zeta(t))} \right|_{t=0},$$

which, by a similar computation, is readily seen to agree with the right-hand side of (5.3).

Secondly, to check that every solution has this form, we analyse the leading order behavior, which, for (5.5), is given by

$$F_n(\mathbf{z}, \bar{\mathbf{z}}) = a_n'(z_0) z_n + \dots, \tag{5.6}$$

where the omitted terms depend on $z_0, \bar{z}_0, \dots, z_{n-1}, \bar{z}_{n-1}$, but not on z_n or $\bar{z}_n$. On the other hand, the leading order terms in the general solution to (5.3) can be analysed as follows. According to (5.3) when $k = n$, we find that F satisfies the Cauchy–Riemann equations $\partial F/\partial \bar{z}_n = 0$, and so F is an analytic function of z_n. Moreover, if we differentiate the remaining equations in (5.3) for $k < n$ with respect to z_n, we find that

$$\frac{\partial^2 F}{\partial \bar{z}_k \, \partial z_n} = 0, \quad k = 0, \dots, n-1.$$

Finally, we differentiate (5.3) with respect to $\bar{z}_n$ and use a simple induction to deduce that

$$\frac{\partial^2 F}{\partial z_k \, \partial z_n} = 0, \quad k = 1, \dots, n.$$

Together, these imply that the general solution to (5.3) has the leading order terms

$$F_n(\mathbf{z}, \bar{\mathbf{z}}) = b(z_0) z_n + \dots, \tag{5.7}$$

where b is an analytic function of z_0 and the omitted terms depend on $z_0, \bar{z}_0, \dots, z_{n-1}, \bar{z}_{n-1}$. Thus, using convexity of the domain, we can set $a_n' = b$, and subtract off the solution (5.5) corresponding to a_n to lead to a solution depending only on the remaining variables $z_0, \bar{z}_0, \dots, z_{n-1}, \bar{z}_{n-1}$. Induction, as usual, completes the proof. □

One point of interest is that, although we only need to assume that the functions f, g are C^2, the system (1.1), (5.1) automatically requires them to be real-analytic functions of $\mathbf{x}$, $\mathbf{y}$.

Example 5.3. Consider the case of an 6×6 real Jordan block with two complex eigenvalues $\pm i$, associated with two irreducible 3×3 complex Jordan blocks. Here

$$M = \begin{pmatrix} \Lambda & I & \\ & \Lambda & I \\ & & \Lambda \end{pmatrix} \quad \text{where} \quad \Lambda = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If we let $F = f + ig$, then according to Lemma 5.1, the real system (1.1) is equivalent to the system of complex partial differential equations

$$\frac{\partial F}{\partial \bar{z}_0} = \frac{1}{2}\frac{\partial \bar{F}}{\partial \bar{z}_1} - \frac{1}{4}\frac{\partial \bar{F}}{\partial \bar{z}_2}, \quad \frac{\partial F}{\partial \bar{z}_1} = \frac{1}{2}\frac{\partial \bar{F}}{\partial \bar{z}_2}, \quad \frac{\partial F}{\partial \bar{z}_2} = 0.$$

According to Theorem 5.2, the general solution f, g to (1.1) in this case is a sum

$$f + ig = F = F_0 + F_1 + F_2$$

of the following three "complex solutions":

$$\begin{aligned} F_0(\mathbf{z}) &= a_0(z_0), \\ F_1(\mathbf{z}) &= z_1 a_1'(z_0) + \tfrac{1}{2}\overline{a_1(z_0)}, \\ F_2(\mathbf{z}) &= z_1^2 a_2''(z_0) + 2z_2 a_2'(z_0) + \overline{z_1 a_2'(z_0) - \tfrac{1}{4}a_2(z_0)}, \end{aligned}$$

where a_0, a_1, a_2 are arbitrary analytic functions of the complex variable $z_0 = x_0 + iy_0$.

The next case to consider is when M is reducible, but has just one complex conjugate pair of eigenvalues. The real canonical form of M then consists of several blocks of the form (5.1). The corresponding solution to (1.1) in this case is entirely analogous to the complex case given in Theorem 3.1, but with formulae like those in Theorem 5.2. Thus, corresponding to each block, there is a parametrised complex variable $\zeta^i(t)$, and for each k up to the maximal sized block, one forms parametrised vectors of complex variables $\boldsymbol{\zeta}^{(k)}(t)$, whose entries are determined by the number of blocks of size at least $2k$, cf. (3.3). Finally, the kth solution F_k, as given by (5.2), will have the same expression as in (5.5), but where a_k now depends on the vector $\boldsymbol{\zeta}^{(k)}(t)$. We leave it to the reader to supply the missing details.

Finally, for a general real matrix M, we break it up into macro- and microblocks as in the complex case, each macroblock corresponding either to a single real eigenvalue, or to a complex conjugate pair of eigenvalues. The general solution to (1.1) then decomposes into the sum of individual solutions for each macroblock, cf. (4.2), which are given either by Theorem 3.1 or by the analogous result from the previous paragraph. The statement of this result is then the same as that of Theorem 4.3. Interestingly, we conclude that, for the real case of (1.1), the system implies that f, g are analytic functions of all variables except the major variables corresponding to the Jordan blocks with real eigenvalues.

6. Examples

We now illustrate our general results with a couple of examples arising in applications. We begin with a relatively simple example, which originally arose in the study of conservation laws in planar, anisotropic elasticity [**6**]. Let

$$M = \begin{pmatrix} 0 & \frac{1}{\alpha} & 0 & 0 \\ -1 & 0 & 0 & \beta \\ 0 & 0 & 0 & \alpha \\ 0 & \frac{\beta}{\alpha} & -1 & 0 \end{pmatrix},$$

where the parameters α, β are related to the elastic moduli of an orthotropic elastic material. The characteristic equation for M is

$$\lambda^4 + 2\sigma\lambda^2 + 1 = 0,$$

where

$$\sigma = \frac{\alpha^2 + 1 - \beta^2}{2\alpha}.$$

Strong ellipticity imposes the condition that $\alpha > 0$, $|\beta| < \alpha + 1$, so that $\sigma > 0$, hence the eigenvalues of M are all complex. If $1 - \alpha = \beta$, then $\sigma = 1$, and the elastic material is isotropic, with $\alpha = \mu/(2\mu + \lambda)$ in terms of the standard Lamé moduli. In this case M has a single complex conjugate pair of eigenvalues at $\pm i$, with corresponding eigenvectors

$$\mathbf{a}_\pm = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ \alpha \\ -\alpha \\ 0 \end{pmatrix}.$$

Thus M is irreducible, so we are in the case covered by Theorem 5.2, when $n = 1$. We define the complex variables

$$\xi = (x_1 - x_4) + i(x_2 + x_3), \quad \eta = (x_1 + x_4) + i\alpha(x_2 - x_3).$$

Then, according to Theorem 5.2, the general solution to (1.1) for this particular matrix can be written as

$$f(\mathbf{x}) + ig(\mathbf{x}) = a_0(\eta) + \xi \frac{da_1}{d\eta} + \frac{1 - \alpha}{2} a_1(\eta),$$

where a_0, a_1 are arbitrary complex analytic functions of the variable η.

The other case we consider explicitly is when $\sigma > 1$, so the matrix M has four simple purely imaginary eigenvalues at $\pm \tau i$, $\pm \tau^{-1} i$, where

$$\tau = \sqrt{\sigma + \sqrt{\sigma^2 - 1}} > 1,$$

with corresponding complex eigenvectors

$$\mathbf{a}_\pm = \begin{pmatrix} \beta\tau \\ 0 \\ 0 \\ \tau - \alpha\tau^3 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ \alpha\beta\tau^2 \\ \alpha^2\tau^2 - \alpha \\ 0 \end{pmatrix}, \quad \mathbf{b}_\pm = \begin{pmatrix} \beta\tau^2 \\ 0 \\ 0 \\ \tau^2 - \alpha \end{pmatrix} \pm i \begin{pmatrix} 0 \\ \alpha\beta\tau \\ \alpha^2\tau - \alpha\tau^3 \\ 0 \end{pmatrix}.$$

As a direct consequence of Theorem 5.2, we deduce that every solution f, g to the basic equation (1.1) for this particular matrix M has the form

$$f(\mathbf{x}) = a_1 + b_1, \quad g(\mathbf{x}) = \frac{1}{\tau} a_2 + \tau b_2,$$

where

$$a_1 + ia_2 = a(\xi) \quad \text{and} \quad b_1 + ib_2 = b(\eta)$$

are arbitrary analytic functions of their complex arguments

$$\xi = \mathbf{a}_+ \cdot \mathbf{x}, \quad \eta = \mathbf{b}_+ \cdot \mathbf{x}.$$

This recovers the results in [**4**, Theorem 4.1] on conservation laws in linear isotropic elasticity, and [**6**] in the anisotropic case.

We next consider a more complicated example from the theory of biHamiltonian systems [**7, 9**]. A system of first order ordinary differential equations is called a biHamiltonian system, cf. [**5**, chapter 7], if it can be written in Hamiltonian form in two different ways. Here we look at the case when both Hamiltonian structures are symplectic, and simultaneously linearisable. The system takes the form

$$\dot{x} = J_1 \nabla H_1 = J_2 \nabla H_2, \tag{6.1}$$

where J_1, J_2 are nonsingular invertible matrices, which determine the two Hamiltonian structures, and $H_1(x)$, $H_2(x)$ are the corresponding Hamiltonian functions. Given the two Hamiltonian matrices, then the classification of all corresponding biHamiltonian systems reduces to our basic equation (1.1), i.e. we must solve

$$\nabla H_1 = M \nabla H_2, \tag{6.2}$$

where

$$M = J_1^{-1} . J_2. \tag{6.3}$$

The classification of pairs of Hamiltonian matrices reduces to the classification of skew-symmetric matrix pencils, a problem solved by Weierstrass in the nonsingular case, and Kronecker in the singular case, cf. [**10**]. The general case decomposes into a direct sum of *irreducible* Hamiltonian pairs, corresponding to the Jordan block decomposition of the matrix M. Owing to the skew-symmetry of the matrices J, each eigenvalue and associated Jordan block(s) of M always appear twice. Here we look at the case of a single irreducible, nondegenerate pair; the general reducible, nondegenerate case is discussed in [**7**]. Under a complex change of basis, we can place the Hamiltonian pair in canonical form

$$J_1 = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & -\lambda I - U^T \\ \lambda I + U & 0 \end{pmatrix}, \tag{6.4}$$

where the notation is the same as in (2.1). In particular, J_1 is in the well-known canonical matrix from classical mechanics. We write $\mathbf{x} = (\mathbf{p}, \mathbf{q}) = (p_0, \ldots, p_n, q_0, \ldots, q_n)$ according to the matrix block decomposition (6.4), so that the p's and the q's are the canonically conjugate variables for the standard symplectic structure on $\mathbb{R}^{2n}$ as determined by the matrix J_1.

The matrix M, (6.3), is given by

$$M = \begin{pmatrix} \lambda I + U & 0 \\ 0 & -\lambda I - U^T \end{pmatrix}, \tag{6.5}$$

which would be in Jordan canonical form if we were to reorder the q's in reverse order. Therefore, we are in the case of a single eigenvalue, having two associated Jordan blocks. Theorem 3.1 immediately implies the following:

THEOREM 6.1. *Let $H_1(\mathbf{x})$, $H_2(\mathbf{x})$ be smooth functions on a convex domain, which satisfy* (6.2) *where M is given by* (6.5). *Then there exist smooth scalar-valued functions $a_k(s_1, s_2)$, $k = 0, \ldots, n$, such that*

$$H_1(\mathbf{x}) = H_1^0(\mathbf{x}) + H_1^1(\mathbf{x}) + \ldots + H_1^n(\mathbf{x}),$$
$$H_2(\mathbf{x}) = H_2^0(\mathbf{x}) + H_2^1(\mathbf{x}) + \ldots + H_2^n(\mathbf{x}),$$

where

$$H_1^k(\mathbf{x}) = \lambda \frac{\partial^k}{\partial t^k} a_k(\pi(t), \varpi(t))\bigg|_{t=0} + k \frac{\partial^{k-1}}{\partial t^{k-1}} a_k(\pi(t), \varpi(t))\bigg|_{t=0}, \tag{6.6}$$

$$H_2^k(\mathbf{x}) = \frac{\partial^k}{\partial t^k} a_k(\pi(t), \varpi(t))\bigg|_{t=0}, \tag{6.7}$$

where

$$\pi(t) = p_0 + tp_1 + t^2 p_2 + \ldots + t^n p_n,$$
$$\varpi(t) = q_n + tq_{n-1} + t^2 q_{n-2} + \ldots + t^n q_0.$$

An important consequence of this result is the complete integrability of any irreducible biHamiltonian system (6.1). To see this, consider first the subsystem governing the time evolution of the major variables, which are (p_0, q_n). A calculation shows that these are

$$\begin{aligned} \frac{dp_0}{dt} &= -\frac{\partial H_1}{\partial q_0} = -\lambda\, n!\, \frac{\partial a_n}{\partial q_n}, \\ \frac{dq_n}{dt} &= \frac{\partial H_1}{\partial p_n} = \lambda\, n!\, \frac{\partial a_n}{\partial p_0}. \end{aligned} \tag{6.8}$$

This is just an autonomous two-dimensional (one degree of freedom) Hamiltonian system, with Hamiltonian function $\lambda\, n!\, a_n(p_0, q_n)$, and is easily integrated by quadrature, cf. [**5**]. Thus we can determine the time evolution of (p_0, q_n) explicitly. (Curiously, the canonically conjugate variables p_0, q_0 for the standard symplectic structure determined by J_1 are *not* the canonically conjugate variables p_0, q_n for the reduced system (6.8), nor do they coincide with canonically conjugate variables for the second symplectic structure determined by J_2.)

We now show how the time evolution of the minor variables can also be determined by successively solving a hierarchy of two-dimensional forced linear Hamiltonian systems in the conjugate pairs p_k, q_{n-k}. Consider first p_1, q_{n-1}. A straightforward computation using (3.9) shows that their time evolution is governed by the system

$$\frac{dp_1}{dt} = -\frac{\partial H_1}{\partial q_{n-1}} = -\lambda(n-1)!\, \frac{\partial a_{n-1}}{\partial q_n} - n!\left(\frac{\partial^2 a_n}{\partial p_0\, \partial q_n} p_1 + \frac{\partial^2 a_n}{\partial q_n^2} q_{n-1}\right),$$

$$\frac{dq_{n-1}}{dt} = \frac{\partial H_1}{\partial p_1} = \lambda(n-1)!\, \frac{\partial a_{n-1}}{\partial p_0} + n!\left(\frac{\partial^2 a_n}{\partial p_0^2} p_1 + \frac{\partial^2 a_n}{\partial p_0\, \partial q_n} q_{n-1}\right).$$

Substituting the known solution $p_0(t)$, $q_n(t)$ to (6.8) reduces this to a forced, linear, non-autonomous two-dimensional Hamiltonian system. Integrating this linear system allows us to determine the time evolution of (p_1, q_{n-1}) explicitly. It is not hard to see that this recursive procedure continues, so at the kth stage, to determine the time evolution of (p_k, q_{n-k}) we need only solve a time-dependent two-dimensional linear Hamiltonian system

$$\frac{dp_k}{dt} = -n!\left(\frac{\partial^2 a_n}{\partial p_0\,\partial q_n} p_k + \frac{\partial^2 a_n}{\partial q_n^2} q_{n-k}\right) - G_k(t),$$
$$\frac{dq_{n-k}}{dt} = n!\left(\frac{\partial^2 a_n}{\partial p_0^2} p_k + \frac{\partial^2 a_n}{\partial p_0\,\partial q_n} q_{n-k}\right) + \tilde{G}_k(t),$$

where G_k, $\tilde{G}_k$ are certain explicit functions of $(p_0, \ldots, p_{k-1}, q_n, \ldots, q_{n-k-1})$, whose time evolution has thus already been determined. Therefore, we have explicitly demonstrated that any biHamiltonian system relative to the irreducible, constant eigenvalue Hamiltonian pair can be integrated by solving a single two-dimensional (one degree of freedom) autonomous Hamiltonian system, along with n forced linear, nonautonomous two-dimensional Hamiltonian systems. In this way, we can call such a biHamiltonian system completely integrable.

7. The issue of smoothness

In case all the eigenvalues of M are real, the change of variables that brings M to Jordan form is real, so the arguments of Theorems 2.1 and 3.1 go through without change—all that is needed is for the partial differentiations in each variable to commute with each other, and for a Poincaré lemma to hold; see [**8**, chapter II, §6, Théorème VI and chapter IX, §3, Théorème I].

Here is an example, in which $H(x)$ denotes the Heaviside function, which is 0 for x negative, and 1 for x positive. Consider the 2×2 real matrix

$$M = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Let $g_1(x, y) = yH'(x)$, where the derivative is taken in the distributional sense. From formula (2.5-1) in Section 2, the distributional solution to $\nabla f_1 = M \nabla g_1$ is given by

$$f_1 = \lambda y H'(x) + H(x).$$

This can be verified through a straightforward computation using the definition of derivative of a distribution.

A question of John Ball and Richard James [**1**], arising in their work on phase transitions, which was originally mentioned to us by Robert Hardt, contributed to this work. They asked whether the following result (for B a box or a ball) is true. The proof now follows from our basic results.

THEOREM 7.1. *Let $B \subset \mathbb{R}^n$ be a convex open set. Let E and F be measurable subsets of B. Then the following two statements are equivalent:*

(i) *for all test functions φ* (i.e. C^∞ *functions with compact support in* B)

$$\int_E \nabla\varphi(x)\,dx = 0 \quad \textit{if and only if} \quad \int_F \nabla\varphi(x)\,dx = 0;$$

(ii) $$E = F \quad \textit{or} \quad E = B \backslash F.$$

Proof. The fact that (ii) implies (i) is trivial. Conversely, if (i) holds, then we use the following theorem, cf. [**3**, Théorème 1, p. 109].

THEOREM 7.2. *Let* $L, L_1, \ldots, L_n$ *be linear functionals on a vector space* V. *Then* $L \in \operatorname{span}\{L_1, \ldots, L_n\}$ *if and only if* $\ker L \supseteq \ker L_1 \cap \ldots \cap \ker L_n$.

This result implies that there exist real constant matrices M, N such that

$$\nabla\chi_E = M\nabla\chi_F, \quad \text{and} \quad \nabla\chi_F = N\nabla\chi_E,$$

where χ_E denotes the characteristic function of the set E. The proof of the theorem follows by checking the various possibilities for M. If M is a multiple of the identity, it must be $\pm I$, and the result follows. If M has only complex eigenvalues, then E and F must be either empty or all of B since solutions of the system must be real-analytic. If M has some real and some complex eigenvalues, there is a coordinate system in which M has a direct sum decomposition in which one part has all complex eigenvalues, and the other has all real eigenvalues. In this case, they must be of the form (2.5-0) in Section 2, since characteristic functions cannot have non-constant polynomial form. But this gives the desired result.

Acknowledgments

We would like to thank Robert Hardt, John Ball and Dick James for motivating and encouraging this work. We also thank Jesus Gil de Lamadrid for the functional analysis references. One of us (P.J.O.) would like to thank the Institute for Mathematics and its Applications (I.M.A.) for the support during the time this work was done.

References

1 J. M. Ball and R. James. Work in progress.
2 L. Comtet. *Advanced Combinatorics; the Art of Finite and Infinite Expansions* (Boston: D. Reidel, 1974).
3 J. Dieudonné. La dualité dans les espaces vectoriels topologiques. *Ann. Sci. École Norm. Sup. (3)* **59** (1942), 107–139.
4 P. J. Olver. Conservation laws in elasticity. II. Linear homogeneous isotropic elastostatics. *Arch. Rational Mech. Anal.* **85** (1984), 131–160.
5 P. J. Olver. *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics 107 (New York: Springer, 1986).
6 P. J. Olver. Conservation laws in elasticity. III. Planar linear anisotropic elastostatics. *Arch. Rational Mech. Anal.* **102** (1988), 167–181.
7 P. J. Olver. Canonical variables and integrability of biHamiltonian systems *Phys. Lett. A* (to appear).
8 L. Schwartz. *Théorie des Distributions*, nouvelle édition (Paris: Hermann, 1966).
9 F.-J. Turiel. Classification locale d'un couple de formes symplectiques Poisson-compatibles. *Comptes Rendus Acad. Sci. Paris* **308** (1989), 575–578.
10 H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices* (London; Blackie, 1932).
11 F. W. Warner. *Foundations of Differentiable Manifolds and Lie Groups* (Glenview, Ill.: Foresman, 1971).

(*Issued* 14 *December* 1990)